

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-137630

(43)Date of publication of application : 31.05.1996

(51)Int.Cl.

G06F 3/06

G06F 13/10

(21)Application number : 06-277422

(71)Applicant : HITACHI LTD

(22)Date of filing : 11.11.1994

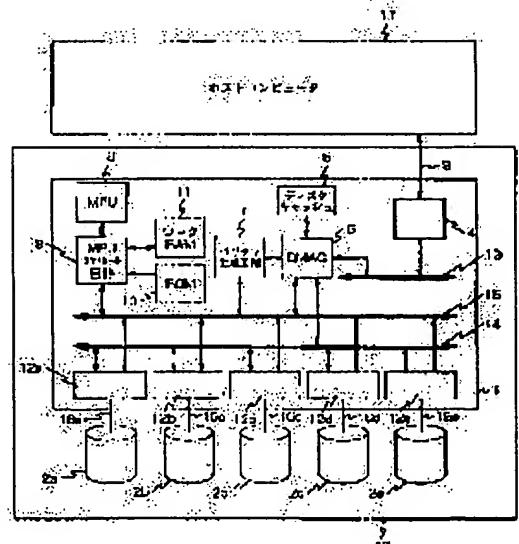
(72)Inventor : ICHIKAWA MASATOSHI
ISONO SOICHI
HONDA KIYOSHI
MATSUMOTO JUN
IWASAKI HIDEHIKO

(54) DISK ARRAY CONTROLLER AND DISK ARRAY DEVICE

(57)Abstract:

PURPOSE: To provide the disk array controller and the disk array device at a low cost in which data are transferred at a high speed.

CONSTITUTION: This controller is provided with a disk array control section having an MPU 8, a user data transfer control section consisting of host interfaces 3, 4 with respect to a host computer 17, a memory 5 storing tentatively data, a redundant data generating circuit 7 generating redundant data, plural channels of disk device interfaces 16a-16e, 12a-12e, a host interface, a memory 5 and a data transfer control means (DMAC) 6 controlling the data transfer among the host interface, the memory 5, the redundant data generating circuit 7 and the disk interface, and at least three-bus configuration is adopted for buses for a control (MPU) bus 15, a host data bus 13 and a drive data bus 14.



LEGAL STATUS

[Date of request for examination] 12.11.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application
converted registration]

[Date of final disposal for application]

[Patent number] 3581727

[Date of registration] 30.07.2004

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

JPO and NCIPPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] The host interface [it has the disk array control section which has one or more MPU, and] between host computers, The memory which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said memory, said redundancy data generation means, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation means, and said data transfer control means of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host interfaces and said memory by said data transfer control means in said user data transfer control section, The disk array controller characterized by having the drive data bus which performs data transfer between said disk unit interfaces and said memory by said data transfer control means in said user data transfer control section.

[Claim 2] The disk array controller according to claim 1 characterized by constituting said host interface from a SCSI interface.

[Claim 3] The disk array controller according to claim 1 characterized by constituting said disk unit interface from a SCSI interface.

[Claim 4] Said data transfer control means is a disk array controller according to claim 1 characterized by constituting data transfer pass possible [two or more assignment] between the field of said memory, and said disk unit interface.

[Claim 5] the disk array controller according to claim 1 characterized by constituting said data transfer control means from two or more registers which specify the channel of said disk unit interface corresponding to two or more counters which specify the address of said memory, and the counter of each this.

[Claim 6] The disk array controller according to claim 1 characterized by constituting so that data may be transmitted using two or more transfer pass between the field of said memory, and said redundancy data generation means.

[Claim 7] Said data transfer control means is a disk array controller according to claim 1 characterized by constituting by DMA.

[Claim 8] the disk array controller according to claim 1 by which it is constituting [said data-transfer control means has two or more registers which specify the channel of a disk unit interface corresponding to two or more counters which specify the address of said memory, and the counter of each this, and constituted them possible / two or more assignment of data-transfer pass / between the field of said memory, and said disk unit interface, and / so that the data of data-transfer pass may input into said redundancy data generation means corresponding to said data-transfer pass further] characterized.

[Claim 9] The disk array controller according to claim 1 characterized by equipping said disk unit interface with a drive interface controller according to each channel.

[Claim 10] The host interface [it has the disk array control section which has one or more MPU, and] between host computers, The memory which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said memory, said redundancy data generation means, and said disk unit interface. The bus for control which controls the disk unit interface, said redundancy data generation means, and said data transfer control means of said multiple channel of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host

interfaces and said memory by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said memory by said data transfer control means in said user data transfer control section. The disk array controller characterized by having linked said redundancy data generation means with said data transfer control means directly, having installed between said memory and said drive data buses, and mitigating TORAHIIKKU to said memory in redundancy data generation.

[Claim 11] The host interface [it has the disk array control section which has one or more MPU, and] between host computers, The memory which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said memory, said redundancy data generation means, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation means, and said data transfer control means of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host interfaces and said memory by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said memory by said data transfer control means in said user data transfer control section. Said data transfer control means constitutes data transfer pass possible [two or more assignment] between the field of said memory, and said disk unit interface. Furthermore, the disk array controller characterized by constituting so that the data of data transfer pass may be inputted into said redundancy data generation means corresponding to this data transfer pass.

[Claim 12] The host interface [it has the disk array control section which has one or more MPU, and] between host computers, The memory which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said memory, said redundancy data generation means, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation means, and said data transfer control means of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host interfaces and said memory by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said memory by said data transfer control means in said user data transfer control section. The disk array controller characterized by constituting so that the redundancy data to user data may be generated in said redundancy data generation means and it may transmit to said disk unit interface.

[Claim 13] The host interface [it has the disk array control section which has one or more MPU, and] between host computers, The memory which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said memory, said redundancy data generation means, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation means, and said data transfer control means of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host interfaces and said memory by said data transfer control means in said user data transfer control section, The disk array controller which has the drive data bus which performs data transfer between said disk unit interfaces and said memory by said data transfer control means in said user data transfer control section is formed. Disk array equipment characterized by preparing two or more disk units connected to said disk unit interface of this disk array controller.

[Claim 14] The host interface [it has the disk array control section which has one or more MPU, and] between host computers, The memory which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said memory, said redundancy data generation means, and said disk unit interface. The bus for control which controls the disk unit interface, said redundancy data generation means, and said data transfer control means of said multiple channel of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host

interfaces and said memory by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said memory by said data transfer control means in said user data transfer control section. The disk array controller which linked said redundancy data generation means with said data transfer control means directly, installed between said memory and said drive data buses, and mitigated TORAIKKU to said memory in redundancy data generation is formed. Disk array equipment characterized by preparing two or more disk units connected to said disk unit interface of this disk array controller.

[Claim 15] The host interface [it has the disk array control section which has one or more MPU, and] between host computers, The memory which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said memory, said redundancy data generation means, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation means, and said data transfer control means of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host interfaces and said memory by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said memory by said data transfer control means in said user data transfer control section. Said data transfer control means constitutes data transfer pass possible [two or more assignment] between the field of said memory, and said disk unit interface. Furthermore, the disk array controller constituted so that the data of data transfer pass might be inputted into said redundancy data generation means corresponding to this data transfer pass is formed. Disk array equipment characterized by preparing two or more disk units connected to said disk unit interface of this disk array controller.

[Claim 16] The host interface [it has the disk array control section which has one or more MPU, and] between host computers, The memory which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said memory, said redundancy data generation means, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation means, and said data transfer control means of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host interfaces and said memory by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said memory by said data transfer control means in said user data transfer control section. The disk array controller constituted so that the redundancy data to user data might be generated in said redundancy data generation means and it might transmit to said disk unit interface is formed. Disk array equipment characterized by preparing two or more disk units connected to said disk unit interface of this disk array controller.

[Claim 17] Disk array equipment according to claim 13, 14, 15, or 16 characterized by constituting said data transfer control means in said disk array controller by DMAC.

[Claim 18] Disk array equipment according to claim 13, 14, 15, or 16 characterized by constituting said host interface from a SCSI interface.

[Translation done.]

* NOTICES *

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to the disk array equipment and the disk array controller of architecture of RAID5 grade.

[0002]

[Description of the Prior Art] In the architecture of the conventional disk array controller, the number of internal buses was one, or it was two, the bus for control by internal organs MPU, and the bus for user data transfer during a host interface-disk unit interface.

[0003] moreover, JP,6-180623,A -- coming -- the time -- carrying out -- having -- **** -- as -- a host interface and a disk unit interface -- specialized circuits, such as FIFO (First InFirst Out), were required for speed regulation in between.

[0004]

[The technical problem which invention will solve and to carry out] In the above-mentioned conventional technique, when a host interface and a disk unit interface were SCSI, the transfer rate of the internal bus of a disk array controller and bus width of face had the technical problem that it was restricted by the SCSI control LSI to be used.

[0005] On the other hand, in order to have increased the transfer rate of the internal bus of a disk array controller, and bus width of face for improvement in the speed of disk array equipment, without being dependent on the SCSI control LSI, specialized circuits, such as FIFO, were needed between the SCSI interface and the internal bus, and it had become the factor of the formation of an expensive rank of a disk array controller and disk array equipment by the disk array controller with a host interface and two or more disk unit interfaces.

[0006] That the above-mentioned technical problem should be solved, the purpose of this invention is a low price and is to offer the disk array controller which timed improvement in the speed of data transfer, and its disk array equipment.

[0007] Moreover, other purposes of this invention reduce TORAHIKKU to memory (disk cache) to a parity data generate time, and are to offer the disk array controller which aimed at improvement in an effective transfer rate, and its disk array equipment.

[0008] Moreover, other purposes of this invention are to offer the disk array controller which enabled the data aggregate and distribution, and parity data generation, and its disk array equipment.

[0009]

[Means for Solving the Problem] The host interface [this invention is equipped with the disk array control section which has one or more MPU in order to attain the above-mentioned purpose, and] between host computers, The memory (disk cache) which carries out the temporary storage of the data, and a redundancy data generation means to generate redundancy data (redundancy data generation circuit), It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said disk cache, said redundancy data generation circuit, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation circuit, and said data transfer control means of said user data transfer control section by this disk array control section (MPU bus), The host data bus which performs data transfer between said host interfaces and said disk caches by said data transfer control means in said user data transfer control section, In said user data transfer control section by said data transfer control means It is the

disk array controller or disk array equipment characterized by having the drive data bus which performs data transfer between said disk unit interfaces and said disk caches. Moreover, this invention is characterized by constituting said host interface and a disk unit interface from SCSI (Small Computer System Interface) interfaces, such as SCSI-2, in said disk array controller or disk array equipment. Moreover, this invention is characterized by said data transfer control means constituting data transfer pass possible [two or more assignment] between the field of said disk cache, and said disk unit interface in said disk array controller or disk array equipment. moreover, this invention is characterized by constituting said data transfer control means from two or more registers which specify the channel of said disk unit interface corresponding to two or more counters which specify the address of said disk cache, and the counter of each this in said disk array controller or disk array equipment. Moreover, this invention is characterized by constituting said data transfer control means so that data may be transmitted using two or more transfer pass between the field of said disk cache, and said redundancy data generation circuit in said disk array controller or disk array equipment. Moreover, this invention is set to said disk array controller or disk array equipment, and said data transfer control means is DMAC (Direct Memory Access Controller: part which controls the data transfer between memory <--> memory or between memory <-->I/O instead of CPU or an I/O processor in a small computer.). A source address and the destination address required for data transfer are generated, and the read-out cycle of the source and the write cycle of a destination are driven. It is characterized by constituting. This invention is set to said disk array controller or disk array equipment. Moreover, said data transfer control means It has two or more registers which specify the channel of a disk unit interface corresponding to two or more counters which specify the address of said disk cache, and the counter of each this. It considers as the constituting [constituted possible / two or more assignment of data transfer pass / between the field of said disk cache, and said disk unit interface, and] description so that the data of data transfer pass might be further inputted into said redundancy data generation circuit corresponding to said data transfer pass. Moreover, this invention is characterized by equipping said disk unit interface with a drive interface controller according to each channel in said disk array controller or disk array equipment.

[0010] Moreover, the host interface [this invention is equipped with the disk array control section which has one or more MPU, and] between host computers, The disk cache which carries out the temporary storage of the data, and the redundancy data generation circuit which generates redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said disk cache, said redundancy data generation circuit, and said disk unit interface. The bus for control which controls the disk unit interface, said redundancy data generation circuit, and said data transfer control means of said multiple channel of said user data transfer control section by this disk array control section (MPU bus), The host data bus which performs data transfer between said host interfaces and said disk caches by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said disk caches by said data transfer control means in said user data transfer control section. Said redundancy data generation circuit is directly linked with said data transfer control means. Between said disk caches and said drive data buses It is the disk array controller or disk array equipment characterized by having installed and mitigating TORAHIIKKU to said disk cache in redundancy data generation.

[0011] Moreover, the host interface [this invention is equipped with the disk array control section which has one or more MPU, and] between host computers, The disk cache which carries out the temporary storage of the data, and the redundancy data generation circuit which generates redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said disk cache, said redundancy data generation circuit, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation circuit, and said data transfer control means of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host interfaces and said disk caches by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said disk caches by said data transfer control means in said user data transfer control section. Said data transfer control means constitutes data transfer pass possible [two or more assignment] between the field of said disk cache, and said disk unit interface. Furthermore, it is the disk array controller or disk array equipment characterized by constituting so that the data of data transfer pass may be inputted into said redundancy data generation circuit corresponding to this data transfer pass.

[0012] Moreover, the host interface [this invention is equipped with the disk array control section which has one or more MPU, and] between host computers, The disk cache which carries out the temporary storage of the data, and the redundancy data generation circuit which generates redundancy data, It has the disk unit interface of a multiple channel, and one or more channels. It has the user data transfer control section which consisted of data transfer control means which control the data transfer between said host interface, said disk cache, said redundancy data generation circuit, and said disk unit interface. The bus for control which controls said disk unit interface, said redundancy data generation circuit, and said data transfer control means of said user data transfer control section by this disk array control section, The host data bus which performs data transfer between said host interfaces and said disk caches by said data transfer control means in said user data transfer control section, It has the drive data bus which performs data transfer between said disk unit interfaces and said disk caches by said data transfer control means in said user data transfer control section. It is the disk array controller or disk array equipment characterized by constituting so that the redundancy data to user data may be generated in said redundancy data generation circuit and it may transmit to said disk unit interface.

[0013] This invention one or more Namely, the host interface of a channel, a disk cache, The user data transfer control section which consist of a disk unit interface of DMAC of a channel, a redundancy data generation circuit, and a multiple channel, The bus for control have the disk array control section which has one or more MPU, and according to MPU further (MPU bus), The host data bus which performs data transfer between said host interfaces and said disk caches, It is the disk array controller characterized by having the drive data bus which performs data transfer between said disk unit interfaces and said disk caches.

[0014]

[Function] By having made it the configuration of the bus for control (MPU bus), a host data bus, and a drive data bus in the disk array controller, by said configuration, even if it used the cheap SCSI interface for the host interface and the disk unit interface, the utilization factor of an internal bus is reduced, improvement in the transfer rate of an internal bus is aimed at, as a result, it is cheap, and the disk array equipment with which data transfer was accelerated can be realized.

[0015] Moreover, the data aggregate and distribution are attained by the transfer using [on between the field on a disk cache, and disk unit interfaces, and] two or more pass by said configuration.

[0016] Moreover, in between the field on a disk cache, and parity generation circuits, redundancy data generation (parity data generation) is attained by said configuration using the data stored in a disk cache by the transfer using two or more pass. Moreover, the redundancy data generation circuit (parity data generation circuit) is directly linked with the data transfer control means (DMAC). It makes it possible to input the data of data transfer pass also into a parity generation circuit corresponding to the data transfer pass between the field on a disk cache, and a disk unit interface. The data on a disk cache are transmitted to both a disk unit interface and a parity generation circuit. The generated redundancy data (parity data) can be transmitted to a direct disk unit, without transmitting to a disk cache. The traffic to the disk cache in a redundancy data generate time (parity data generate time) is reduced, and the effective transfer rate in disk array equipment can be raised to the light instruction to the field which continued with large capacity.

[0017]

[Example] The example of this invention is concretely explained using drawing 1 - drawing 5 . Drawing 1 shows the structure-of-a-system Fig. which consists of a host computer 17 which is one example of this invention, and disk array equipment 18. A host computer 17 and disk array equipment 18 are connected with the host interface 3. A host interface 3 is SCSI-2 of the 2-byte bus width of face whose maximum transfer rates it is one channel and are 20 MB/s. In addition, a host interface 3 may consist of two or more channels. A host computer 17 minds a host interface 3, and is SCSI (Small Computer System Interface: a kind of the peripheral-device interface for small computers.). Specification is carried out by ANSI. Disk array equipment 18 is accessed according to a protocol.

[0018] Disk array equipment 18 consists of an array controller 1 and two or more disk units 2. The architecture of RAID5 grade is used for disk array equipment 18, RAID5 redundancy is one parity disk to four data disks, and it makes five disk units the unit of the train configuration of RAID5. Moreover, the disk array controller 1 had the drive interface 16 of five channels, and has connected two or more disk units 2. The drive interface 16 is SCSI-2 of the 1-byte bus width of face whose maximum transfer rates are 10 MB/s.

[0019] Furthermore, the engine performance of the disk unit 2 used with disk array equipment 18 is shown in drawing 2 . That is, for the spindle engine speed in a disk unit, 5400rpm and the average of seek time are [7.2 MB/s and the SCSI data transfer rate of a sustain data transfer rate] 10MB/s for 9ms.

[0020] Next, the configuration of the disk array controller 1 is explained. The array controller 1 is the disk cache (disk cache) 5 which consisted of DRAMs which carry out the temporary storage of the data to MPU8, the MPU control circuit 9, ROM10, the disk array control section that consists of a work piece RAM 11, and the host interface controller 4 and a high speed, and DMAC (Direct Memory Access Controller: part which controls the data transfer between memory <--> memory or between memory <-->I/O instead of CPU or an I/O processor in a small computer.) which is a data transfer control means. A source address and the destination address required for data transfer are generated, and the read-out cycle of the source and the write cycle of a destination are driven. It has three buses of 6, the user data transfer control section which consists of every one drive interface controllers 12a-12e corresponding to the drive interface of 7 or 5 parity generation circuits (redundancy data generation circuit) and the host data bus 13, the drive data bus 14, and the MPU bus 15. MPU8 controls DMAC6 which constitutes the user data control section through the MPU bus 15, the parity generation circuit 7, and the drive interface controllers 12a-12e by the MPU control circuit 9, writing in, reading and making it a work piece RAM 11 based on the program memorized by ROM10. DMAC6 controls the user data transfer on the host data bus 13 and the drive data bus 14, and performs the data aggregate and distribution peculiar to RAID5. The parity generation circuit 7 calculates the parity data which are redundancy data to user data, and one set of a disk unit 2 performs recovery of user data at the time of a failure. The parity generation circuit 7 is directly linked with DMAC6, and exists between a disk cache 5 and the drive data bus 14.

[0021] The host data bus 13 is a bus for user data transfer between the host interface controller 4 and a disk cache 5. The drive data bus 14 is a bus for user data transfer between a disk cache 5 and the drive interface controllers 12a-12e. The MPU bus 15 is a bus for control which controls DMAC6 which constitutes the user data control section, the parity generation circuit 7, and the drive interface controllers 12a-12e.

[0022] Thus, by having made it the configuration of three buses 13, 14, and 15, it was cheap to the host interface controller 4 and the drive interface controller 12, and the SCSI control LSI whose maximum transfer rates are 20 MB/s was able to be adopted as them. That is, the data transfer between a host 17 and a disk cache 5 and the data transfer between a disk cache 5 and disk units 2a-2e can be performed to coincidence, without needing specialized circuits, such as FIFO (First In First Out), by having made it the configuration of three buses 13, 14, and 15 in addition to the SCSI control LSI between a host interface 3, the drive interface 16 between the host data buses 13, and the drive data bus 14. The maximum transfer rate of SCSI-2 of the 2-byte bus width of face in a host 17 3, i.e., the host interface which is an interface between users Since it is 20MB/s, it becomes 20 MB/s from specification, respectively about the transfer rate of the host data bus 13 and the drive data bus 14. The data transfer rate of a disk cache 5 In order to perform the transfer between a host 17 and a disk cache 5, and the transfer between a disk cache 5 and disk units 2a-2e to coincidence, it becomes 40 MB/s and data transfer becomes sufficiently possible.

[0023] DMAC6 has three channels, Channel a, Channel b, and Channel c, and the coincidence transfer using three channels is possible for it. That is, Channel a is used for the transfer between the disk cache 5-disk units 2a-2e, the transfer between the disk cache 5-parity generation circuit 7 and a disk cache 5, and the transfer between the disk cache 5-parity generation circuit 7-disk units 2a-2e. Channel b and Channel c are only used for the data transfer between [for / 2a-2e / disk cache 5-disk units] and between the host 17-disk caches 5.

[0024] Next, the function of Channel a in DMAC6 and the data transfer by the channel a in DMAC6 are explained. That is, in DMAC6, it has five port assignment registers 202 corresponding to five cache address counters 201 and cache address counters 201, and five concurrency transfer assignment registers 205 corresponding to the cache address counter 201 about DMAC6 channel a. The cache address counter 201 is a counter which specifies the address of the field 203 on a disk cache 5. The port assignment register 202 is a register which specifies the channel or the parity generation circuit 7 of the drive interfaces 16a-16e corresponding to the cache address counter 201. It is possible to specify five pass 204 between the disk cache 5-drive interfaces 16a-16e or between the disk cache 5-parity generation circuits 7 with the port assignment register 202 corresponding to the cache address counter 201 and this. The channel a in DMAC6 is time sharing, and has the function to transmit the data of a unit with which two or more specified transfer pass 204 became settled in order. Moreover, the concurrency transfer assignment register 205 is a register which specifies carrying out the concurrency transfer of the data of the corresponding pass 204 also in the parity generation circuit 7. In addition to a transfer facility with two or more pass 204, the channel a in DMAC6 has the function which carries out the concurrency transfer of the data of two or more pass 204 also in the parity generation circuit 7.

[0025] Drawing 3 shows the case where the channel a in DMAC6 is used for the data transfer between the disk

cache 5-disk units 2a-2e. Drawing 4 shows the case where the channel a in DMAC6 is used for the data transfer between the disk cache 5-parity generation circuit 7-disk caches 5. Drawing 5 shows the case where the channel a in DMAC6 is used for the transfer between the disk cache 5-disk units 2a-2d, and the coincidence parallel transfer between disk cache 5-parity generation circuit 7-disk unit 2e.

[0026] As shown in drawing 3, MPU8 through the MPU bus 15 to the cache address registers 201a-201e in the channel a of DMAC6 The address of the fields 203a-203e of a disk cache 5 is specified. By specifying the channel of the drive interfaces 16a-16e as the port assignment registers 202a-202e corresponding to the cache address registers 201a-201e Five pass between the fields 203a-203e on a disk cache 5 and the drive interfaces 16a-16e can be specified. Moreover, the disk unit 2 which performs data transfer is beforehand chosen by SCSI command issue, and the pass of the drive interface 16a - 16e-disk units 2a-2e exists for every channel of the drive interfaces 16a-16e.

[0027] Therefore, in DMAC6, if a transfer facility with two or more pass 204a-204e of Channel a is used, the user data stored on the disk cache 5 using the host data bus 13 from the host 17 are divided into the fixed magnitude called stripe size, and the distributed transfer to two or more disk units 2a-2e is possible. Moreover, even when the above-mentioned user data are stored in two or more discontinuity fields on a disk cache 5, the distributed transfer to disk units 2a-2e which are different, respectively in the data of the field 203 of 5 of the arbitration on a disk cache 5 is possible.

[0028] Moreover, it is possible to carry out the set transfer of the data from two or more disk units 2a-2e to the continuation field on a disk cache 5 by the transfer facility with these two or more pass by changing the direction of a data transfer from a disk unit 2 to a host 17.

[0029] Furthermore, as shown in drawing 4, it sets to the channel a in DMAC6. By specifying the parity generation circuit 7 as the port assignment registers 202a-202d, and specifying two or more pass 204a-204d between two or more fields 203a-203d on a disk cache 5, and the parity generation circuit 7 The parity data operation in the parity generation circuit 7 using user data [in / by the transfer facility based on specified this two or more pass 204a-204d / two or more fields 203a-203d on a disk cache 5], It becomes possible to transmit the calculated this parity data to field 203e of a disk cache 5 through port assignment register 202e and cache address register 201e.

[0030] moreover, the data transfer according [the channel / in / as shown in drawing 5 / DMAC6 / a] to the above-mentioned two or more pass 204a-204d -- in addition, it has the function which carries out the parallel input of the each pass [204a-204d] transfer data in the parity generation circuit 7 with the parallel transfer assignment registers 205a-205d corresponding to each pass 204a-204d. The remaining pass 204e of the five lines is used. Moreover, the output of the parity generation circuit 7 In the channel [in / the result / it can store in disk unit 2e through cache address register 201e, port assignment register 202e, and parallel transfer assignment register 205e, and / DMAC6] a The data transfer of disk cache 5-disk unit 2a - a between and the coincidence parallel data transfer between disk cache 5-parity generation circuit 7-disk unit 2e become possible.

[0031] With disk array equipment 18, about the lead instruction to a continuation field, the channel a in DMAC6 and Channel b are used, and data transfer between the disk unit 2a - 2e-disk caches 5 and data transfer between the host 17-disk caches 5 are performed to coincidence with the large capacity from a host 17. The set transfer of the data is carried out to the continuation field on a disk cache 5 at the data transfer between the disk unit 2a - 2e-disk caches 5 using a transfer facility with two or more pass 204a-204e of the channel a in DMAC6 from two or more disk units 2a-2e.

[0032] Moreover, with disk array equipment 18, about the light instruction to a continuation field, the channel a in DMAC6 and Channel b are used, and data transfer between the disk cache 5-disk units 2a-2e and data transfer between the host 17-disk caches 5 are performed to coincidence with the large capacity from a host 17. In the data transfer between the disk cache 5-disk units 2a-2d A transfer facility with two or more pass 204a-204d of the channel a in DMAC6, The function which carries out the parallel input of the transfer data in each pass 204a-204d in the parity generation circuit 7 is used. The data of a disk cache 5 Two or more disk units 2a-2d, The parity data which transmitted to both parity generation circuits 7 at coincidence, and were generated in the parity generation circuit 7 It transmits to disk unit 2e directly through cache address register 201e, port assignment register 202e, and parallel transfer assignment register 205e using pass 204e through a disk cache 5. [0033] In this example, the disk array control section is considered as the single processor configuration. To the host interface controller 4 and the drive interface controller 12 It is cheap and the SCSI control LSI of marketing whose maximum transfer rates are 20 MB/s is adopted. Since specialized circuits, such as FIFO, do

not exist other than the SCSI control LSI between a host interface 3, the drive interface 16 between the host data buses 13, and the drive data bus 14, A disk array control section the cost price of a disk array controller with a multiprocessor configuration (eight MPU) By the configuration of the MPU bus 15 and two internal buses 13 and 14 of a user data transfer bus, it can decrease to one fifth and, as a result, low-pricing can be attained as disk array equipment 18.

[0034] Moreover, since it constituted from disk array equipment 18 so that the parity data generated in the parity generation circuit might be transmitted to a direct disk unit, parity can be raised about 40% compared with the method which once memorizes the effective transfer rate of the light instruction to a continuation field to a disk cache 5 with the large capacity from a host 17.

[0035] Furthermore, since SCSI-2 which continues to be extended as a standard interface is adopted as a drive interface 16, a future highly efficient disk unit is connected and it becomes possible to constitute highly efficient disk array equipment.

[0036]

[Effect of the Invention] According to this invention, by having had the user data transfer control-section capital disk array control section, and having made it at least three bus arrangements, the bus for control (MPU bus), a host data bus, and a drive data bus, use of the cheap interface controller LSI is enabled as a host interface controller and a drive interface controller, and the effectiveness that not only a disk array controller but also the cost price can be sharply reduced also as disk array equipment as a result is done so.

[0037] Moreover, according to this invention, the transfer rate of the internal bus of a disk array controller does so the effectiveness that it cannot become the engine-performance neck of disk array equipment, and the transfer rate of an internal bus can be raised, and reduction of bus width of face can also be aimed at. Moreover, according to this invention, the effectiveness which can improve the effective transfer rate of disk array equipment to the light instruction to the field which reduced the traffic of a disk cache to the parity generate time, and continued with the large capacity from a user in the disk array controller is done so.

[Translation done.]

(51)Int.Cl.⁶G 06 F 3/06
13/10識別記号 540
340 B 7368-5E

F I

技術表示箇所

審査請求 未請求 請求項の数18 O L (全13頁)

(21)出願番号 特願平6-277422
 (22)出願日 平成6年(1994)11月11日

(71)出願人 000005108
 株式会社日立製作所
 東京都千代田区神田駿河台四丁目6番地
 (72)発明者 市川 正敏
 神奈川県川崎市麻生区王禅寺1099株式会社
 日立製作所システム開発研究所内
 (72)発明者 磯野 聰一
 神奈川県川崎市麻生区王禅寺1099株式会社
 日立製作所システム開発研究所内
 (72)発明者 本田 聖志
 神奈川県川崎市麻生区王禅寺1099株式会社
 日立製作所システム開発研究所内
 (74)代理人 弁理士 高橋 明夫 (外1名)

最終頁に続く

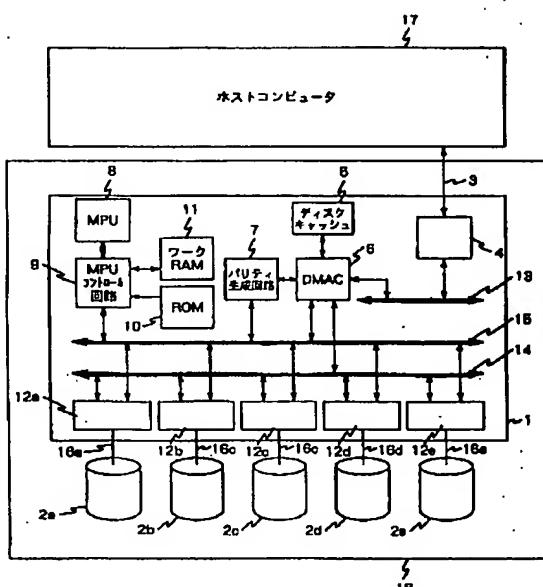
(54)【発明の名称】ディスクアレイコントローラ及びディスクアレイ装置

(57)【要約】

【目的】本発明の目的は、低価格で、データ転送の高速化をはかったディスクアレイコントローラ及びディスクアレイ装置を提供することにある。

【構成】本発明は、MPU 8を有するディスクアレイ制御部を備え、ホストコンピュータ17との間におけるホストインターフェイス3、4と、データを一時格納するメモリ5と、冗長データを生成する冗長データ生成回路7と、複数チャネルのディスク装置インターフェイス16a～16e、12a～12eと、前記ホストインターフェイス、前記メモリ、前記冗長データ生成回路及び前記ディスク装置インターフェイスの間のデータ転送を制御するデータ転送制御手段(DMAC)6とから構成されたユーザデータ転送制御部を備え、内部バスを、少なくとも制御用(MPU)バス15、ホストデータバス13及びドライブデータバス14の少なくとも3バス構成したことを特徴とするものである。

図 1



【特許請求の範囲】

【請求項1】一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインタフェイスと、データを一時格納するメモリと、冗長データを生成する冗長データ生成手段と、複数チャネルのディスク装置インタフェイスと、一つ又は複数のチャネルを有し、前記ホストインタフェイス、前記メモリ、前記冗長データ生成手段及び前記ディスク装置インタフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって前記ユーザデータ転送制御部の前記ディスク装置インタフェイス、前記冗長データ生成手段及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインタフェイスと前記メモリとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インタフェイスと前記メモリとの間のデータ転送を行うドライブデータバスとを有することを特徴とするディスクアレイコントローラ。

【請求項2】前記ホストインタフェイスをSCSIインタフェイスで構成したことを特徴とする請求項1記載のディスクアレイコントローラ。

【請求項3】前記ディスク装置インタフェイスをSCSIインタフェイスで構成したことを特徴とする請求項1記載のディスクアレイコントローラ。

【請求項4】前記データ転送制御手段は、前記メモリの領域と前記ディスク装置インタフェイスとの間においてデータ転送バスを複数指定可能に構成したことを特徴とする請求項1記載のディスクアレイコントローラ。

【請求項5】前記データ転送制御手段は、前記メモリのアドレスを指定する複数のカウンタと、該各々のカウンタに対応して前記ディスク装置インタフェイスのチャネルを指定する複数のレジスタとで構成したことを特徴とする請求項1記載のディスクアレイコントローラ。

【請求項6】前記メモリの領域と前記冗長データ生成手段との間において複数の転送バスを用いてデータを転送するように構成したことを特徴とする請求項1記載のディスクアレイコントローラ。

【請求項7】前記データ転送制御手段は、DMAによって構成したことを特徴とする請求項1記載のディスクアレイコントローラ。

【請求項8】前記データ転送制御手段は、前記メモリのアドレスを指定する複数のカウンタと、該各々のカウンタに対応してディスク装置インタフェイスのチャネルを指定する複数のレジスタとを有し、前記メモリの領域と前記ディスク装置インタフェイスとの間においてデータ転送バスを複数指定可能に構成し、更に前記データ転送バスに対応してデータ転送バスのデータを前記冗長データ生成手段に入力するように構成したこと特徴とする請求項1記載のディスクアレイコントローラ。

【請求項9】前記ディスク装置インタフェイスには、各チャネルに応じてドライブインタフェイスコントローラを備えたことを特徴とする請求項1記載のディスクアレイコントローラ。

【請求項10】一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインタフェイスと、データを一時格納するメモリと、冗長データを生成する冗長データ生成手段と、複数チャネルのディスク装置インタフェイスと、一つ又は複数のチャネルを有し、前記ホストインタフェイス、前記メモリ、前記冗長データ生成手段及び前記ディスク装置インタフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって前記ユーザデータ転送制御部の前記複数チャネルのディスク装置インタフェイス、前記冗長データ生成手段及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインタフェイスと前記メモリとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インタフェイスと前記メモリとの間のデータ転送を行うドライブデータバスとを有し、前記冗長データ生成手段を前記データ転送制御手段に直結して前記メモリと前記ドライブデータバスとの間に設置して冗長データ生成において前記メモリへのトラヒックを軽減したことを特徴とするディスクアレイコントローラ。

【請求項11】一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインタフェイスと、データを一時格納するメモリと、冗長データを生成する冗長データ生成手段と、複数チャネルのディスク装置インタフェイスと、一つ又は複数のチャネルを有し、前記ホストインタフェイス、前記メモリ、前記冗長データ生成手段及び前記ディスク装置インタフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって、前記ユーザデータ転送制御部の前記ディスク装置インタフェイス、前記冗長データ生成手段及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインタフェイスと前記メモリとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インタフェイスと前記メモリとの間のデータ転送を行うドライブデータバスとを有し、前記データ転送制御手段は、前記メモリの領域と前記ディスク装置インタフェイスとの間においてデータ転送バスを複数指定可能に構成し、更に前記データ転送バスに対応してデータ転送バスのデータを前記冗長データ生成手段に入力するように構成したことを特徴とする請求項1記載のディスクアレイコントローラ。

【請求項12】前記データ転送制御手段は、前記メモリのアドレスを指定する複数のカウンタと、該各々のカウンタに対応してディスク装置インタフェイスのチャネルを指定する複数のレジスタとを有し、前記メモリの領域と前記ディスク装置インタフェイスとの間においてデータ転送バスを複数指定可能に構成し、更に前記データ転送バスに対応してデータ転送バスのデータを前記冗長データ生成手段に入力するように構成したことを特徴とする請求項1記載のディスクアレイコントローラ。

に該データ転送バスに対応してデータ転送バスのデータを前記冗長データ生成手段に入力するよう構成したことを特徴とするディスクアレイコントローラ。

【請求項12】一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインタフェイスと、データを一時格納するメモリと、冗長データを生成する冗長データ生成手段と、複数チャネルのディスク装置インタフェイスと、一つ又は複数のチャネルを有し、前記ホストインタフェイス、前記メモリ、前記冗長データ生成手段及び前記ディスク装置インタフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって、前記ユーザデータ転送制御部の前記ディスク装置インタフェイス、前記冗長データ生成手段及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインタフェイスと前記メモリとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インタフェイスと前記メモリとの間のデータ転送を行うドライブデータバスとを有し、前記冗長データ生成手段においてユーザデータに対する冗長データを生成して前記ディスク装置インタフェイスに転送するように構成したことを特徴とするディスクアレイコントローラ。

【請求項13】一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインタフェイスと、データを一時格納するメモリと、冗長データを生成する冗長データ生成手段と、複数チャネルのディスク装置インタフェイスと、一つ又は複数のチャネルを有し、前記ホストインタフェイス、前記メモリ、前記冗長データ生成手段及び前記ディスク装置インタフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって前記ユーザデータ転送制御部の前記ディスク装置インタフェイス、前記冗長データ生成手段及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインタフェイスと前記メモリとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インタフェイスと前記メモリとの間のデータ転送を行うドライブデータバスとを有するディスクアレイコントローラを設け、

該ディスクアレイコントローラの前記ディスク装置インタフェイスに接続された複数のディスク装置を設けたことを特徴とするディスクアレイ装置。

【請求項14】一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけ

るホストインタフェイスと、データを一時格納するメモリと、冗長データを生成する冗長データ生成手段と、複数チャネルのディスク装置インタフェイスと、一つ又は複数のチャネルを有し、前記ホストインタフェイス、前記メモリ、前記冗長データ生成手段及び前記ディスク装置インタフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって前記ユーザデータ転送制御部の前記複数チャネルのディスク装置インタフェイス、前記冗長データ生成手段及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインタフェイスと前記メモリとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インタフェイスと前記メモリとの間のデータ転送を行うドライブデータバスとを有し、前記冗長データ生成手段を前記データ転送制御手段に直結して前記メモリと前記ドライブデータバスとの間に設置して冗長データ生成において前記メモリへのトラヒックを軽減したディスクアレイコントローラを設け、
該ディスクアレイコントローラの前記ディスク装置インタフェイスに接続された複数のディスク装置を設けたことを特徴とするディスクアレイ装置。

【請求項15】一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインタフェイスと、データを一時格納するメモリと、冗長データを生成する冗長データ生成手段と、複数チャネルのディスク装置インタフェイスと、一つ又は複数のチャネルを有し、前記ホストインタフェイス、前記メモリ、前記冗長データ生成手段及び前記ディスク装置インタフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって、前記ユーザデータ転送制御部の前記ディスク装置インタフェイス、前記冗長データ生成手段及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインタフェイスと前記メモリとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インタフェイスと前記メモリとの間のデータ転送を行うドライブデータバスとを有し、前記データ転送制御手段は、前記メモリの領域と前記ディスク装置インタフェイスとの間ににおいてデータ転送バスを複数指定可能に構成し、更に該データ転送バスに対応してデータ転送バスのデータを前記冗長データ生成手段に入力するように構成したディスクアレイコントローラを設け、
該ディスクアレイコントローラの前記ディスク装置インタフェイスに接続された複数のディスク装置を設けたこ

とを特徴とするディスクアレイ装置。

【請求項16】一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインターフェイスと、データを一時格納するメモリと、冗長データを生成する冗長データ生成手段と、複数チャネルのディスク装置インターフェイスと、一つ又は複数のチャネルを有し、前記ホストインターフェイス、前記メモリ、前記冗長データ生成手段及び前記ディスク装置インターフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって、前記ユーザデータ転送制御部の前記ディスク装置インターフェイス、前記冗長データ生成手段及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインターフェイスと前記メモリとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インターフェイスと前記メモリとの間のデータ転送を行うドライブデータバスとを有し、前記冗長データ生成手段においてユーザデータに対する冗長データを生成して前記ディスク装置インターフェイスに転送するように構成したディスクアレイコントローラを設け、

該ディスクアレイコントローラの前記ディスク装置インターフェイスに接続された複数のディスク装置を設けたことを特徴とするディスクアレイ装置。

【請求項17】前記ディスクアレイコントローラにおける前記データ転送制御手段を、DMAによって構成したことを特徴とする請求項13又は14又は15又は16記載のディスクアレイ装置。

【請求項18】前記ホストインターフェイスをSCSIインターフェイスで構成したことを特徴とする請求項13又は14又は15又は16記載のディスクアレイ装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、RAID5等のアーキテクチャのディスクアレイ装置及びディスクアレイコントローラに関するものである。

【0002】

【従来の技術】従来のディスクアレイコントローラのアーキテクチャでは、内部バスが1本であるか、又は内蔵MPUによる制御用バスとホストインターフェイス-ディスク装置インターフェイス間のユーザデータ転送用バスの2本であった。

【0003】また、特開平6-180623号公報に記載されているようにホストインターフェイスとディスク装置インターフェイスの間に速度調整用にFIFO(First InFirst Out)などの専用回路が必要であった。

【0004】

【発明が解決しようする課題】上記従来技術において、

ホストインターフェイス及びディスク装置インターフェイスがSCSIの場合、ディスクアレイコントローラの内部バスの転送速度、バス幅は、使用するSCSIコントロールLSIにより制限されるという課題を有していた。

【0005】一方、ディスクアレイ装置の高速化のためにディスクアレイコントローラの内部バスの転送速度、バス幅を、SCSIコントロールLSIに依存せずに増やすには、SCSIインターフェイスと内部バスの間にFIFOなどの専用回路が必要となり、ホストインターフェイス及びディスク装置インターフェイスを複数チャネルを持つディスクアレイコントローラでは、ディスクアレイコントローラ、ディスクアレイ装置の高価格化の要因となっていた。

【0006】本発明の目的は、上記課題を解決すべく、低価格で、データ転送の高速化をはかったディスクアレイコントローラ及びそのディスクアレイ装置を提供することにある。

【0007】また本発明の他の目的は、パリティデータ生成時に、メモリ(ディスクキャッシュ)へのトラヒックを低減して、実効転送速度の向上をはかったディスクアレイコントローラ及びそのディスクアレイ装置を提供することにある。

【0008】また本発明の他の目的は、データの集合・分散とパリティデータ生成とを可能としたディスクアレイコントローラ及びそのディスクアレイ装置を提供することにある。

【0009】
【課題を解決するための手段】本発明は、上記目的を達成するために、一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインターフェイスと、データを一時格納するメモリ(ディスクキャッシュ)と、冗長データを生成する冗長データ生成手段(冗長データ生成回路)と、複数チャネルのディスク装置インターフェイスと、一つ又は複数のチャネルを有し、前記ホストインターフェイス、前記ディスクキャッシュ、前記冗長データ生成回路及び前記ディスク装置インターフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって前記ユーザデータ転送制御部の前記ディスク装置インターフェイス、前記冗長データ生成回路及び前記データ転送制御手段を制御する制御用バス(MPUバス)と、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインターフェイスと前記ディスクキャッシュとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インターフェイスと前記ディスクキャッシュとの間のデータ転送を行うドライブデータバスとを有することを特徴とするディスクアレイコントローラ又はディスクアレイ装置である。また本発明

は、前記ディスクアレイコントローラ又はディスクアレイ装置において、前記ホストインターフェイス及びディスク装置インターフェイスを、SCSI-2等のSCSI (Small Computer System Interface)インターフェイスで構成したことを特徴とする。また本発明は、前記ディスクアレイコントローラ又はディスクアレイ装置において、前記データ転送制御手段は、前記ディスクキャッシュの領域と前記ディスク装置インターフェイスとの間においてデータ転送バスを複数指定可能に構成したことを特徴とする。また本発明は、前記ディスクアレイコントローラ又はディスクアレイ装置において、前記データ転送制御手段は、前記ディスクキャッシュのアドレスを指定する複数のカウンタと、該各々のカウンタに対応して前記ディスク装置インターフェイスのチャネルを指定する複数のレジスタとで構成したことを特徴とする。また本発明は、前記ディスクアレイコントローラ又はディスクアレイ装置において、前記データ転送制御手段は、前記ディスクキャッシュの領域と前記冗長データ生成回路との間において複数の転送バスを用いてデータを転送するよう構成したことを特徴とする。また本発明は、前記ディスクアレイコントローラ又はディスクアレイ装置において、前記データ転送制御手段は、DMA (Direct Memory Access Controller: 小型コンピュータの中で、CPUやI/Oプロセッサの代わりに、メモリー→メモリー間またはメモリー→I/O間のデータ転送を制御する部分。データ転送に必要な、ソース・アドレス、ディスティネーション・アドレスを生成し、ソースの読み出しサイクル、ディスティネーションの書き込みサイクルを駆動する。)によって構成したことを特徴とする。また本発明は、前記ディスクアレイコントローラ又はディスクアレイ装置において、前記データ転送制御手段は、前記ディスクキャッシュのアドレスを指定する複数のカウンタと、該各々のカウンタに対応してディスク装置インターフェイスのチャネルを指定する複数のレジスタとを有し、前記ディスクキャッシュの領域と前記ディスク装置インターフェイスとの間においてデータ転送バスを複数指定可能に構成し、更に前記データ転送バスに対応してデータ転送バスのデータを前記冗長データ生成回路に入力するよう構成したこと特徴とする。また本発明は、前記ディスクアレイコントローラ又はディスクアレイ装置において、前記ディスク装置インターフェイスには、各チャネルに応じてドライブインターフェイスコントローラを備えたことを特徴とする。

【0010】また本発明は、一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインターフェイスと、データを一時格納するディスクキャッシュと、冗長データを生成する冗長データ生成回路と、複数チャネルのディスク装置インターフェイスと、一つ又は複数のチャネルを有し、前記ホストインターフェイス、前記ディスクキャッシュ、前記

冗長データ生成回路及び前記ディスク装置インターフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって前記ユーザデータ転送制御部の前記複数チャネルのディスク装置インターフェイス、前記冗長データ生成回路及び前記データ転送制御手段を制御する制御用バス (MPUバス)と、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインターフェイスと前記ディスクキャッシュとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インターフェイスと前記ディスクキャッシュとの間のデータ転送を行うドライブデータバスとを有し、前記冗長データ生成回路を前記データ転送制御手段に直結して前記ディスクキャッシュと前記ドライブデータバスとの間に設置して冗長データ生成において前記ディスクキャッシュへのトラヒックを軽減したことを特徴とするディスクアレイコントローラ又はディスクアレイ装置である。

【0011】また本発明は、一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインターフェイスと、データを一時格納するディスクキャッシュと、冗長データを生成する冗長データ生成回路と、複数チャネルのディスク装置インターフェイスと、一つ又は複数のチャネルを有し、前記ホストインターフェイス、前記ディスクキャッシュ、前記冗長データ生成回路及び前記ディスク装置インターフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって、前記ユーザデータ転送制御部の前記ディスク装置インターフェイス、前記冗長データ生成回路及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインターフェイスと前記ディスクキャッシュとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インターフェイスと前記ディスクキャッシュとの間のデータ転送を行うドライブデータバスとを有し、前記データ転送制御手段は、前記ディスクキャッシュの領域と前記ディスク装置インターフェイスとの間においてデータ転送バスを複数指定可能に構成し、更に該データ転送バスに対応してデータ転送バスのデータを前記冗長データ生成回路に入力するよう構成したことを特徴とするディスクアレイコントローラ又はディスクアレイ装置である。

【0012】また本発明は、一つ又は複数のMPUを有するディスクアレイ制御部を備え、ホストコンピュータとの間におけるホストインターフェイスと、データを一時格納するディスクキャッシュと、冗長データを生成する冗長データ生成回路と、複数チャネルのディスク装置インターフェイスと、前記

ンタフェイスと、一つ又は複数のチャネルを有し、前記ホストインターフェイス、前記ディスクキャッシュ、前記冗長データ生成回路及び前記ディスク装置インターフェイスの間のデータ転送を制御するデータ転送制御手段とから構成されたユーザデータ転送制御部を備え、該ディスクアレイ制御部によって、前記ユーザデータ転送制御部の前記ディスク装置インターフェイス、前記冗長データ生成回路及び前記データ転送制御手段を制御する制御用バスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ホストインターフェイスと前記ディスクキャッシュとの間のデータ転送を行うホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インターフェイスと前記ディスクキャッシュとの間のデータ転送を行なうホストデータバスと、前記ユーザデータ転送制御部において前記データ転送制御手段によって前記ディスク装置インターフェイスと前記ディスクキャッシュとの間のデータ転送を行なうホストデータバスと、前記ユーザデータ転送制御部と、一つ又は複数のMPUを有するディスクアレイ制御部を有し、更にMPUによる制御用バス(MPUバス)と、前記ホストインターフェイスと前記ディスクキャッシュとの間のデータ転送を行なうホストデータバスと、前記ディスク装置インターフェイスと前記ディスクキャッシュとの間のデータ転送を行なうドライブデータバスを有することを特徴とするディスクアレイコントローラ又はディスクアレイ装置である。

【0013】即ち、本発明は、一つ又は複数チャネルのホストインターフェイス、ディスクキャッシュ、一つ又は複数チャネルのDMAC、冗長データ生成回路、複数チャネルのディスク装置インターフェイスからなるユーザデータ転送制御部と、一つ又は複数のMPUを有するディスクアレイ制御部を有し、更にMPUによる制御用バス(MPUバス)と、前記ホストインターフェイスと前記ディスクキャッシュとの間のデータ転送を行なうホストデータバスと、前記ディスク装置インターフェイスと前記ディスクキャッシュとの間のデータ転送を行なうドライブデータバスを有することを特徴とするディスクアレイコントローラである。

【0014】

【作用】前記構成により、ディスクアレイコントローラにおいて、ホストインターフェイス及びディスク装置インターフェイスに、安価なSCSIインターフェイスを使用しても、制御用バス(MPUバス)、ホストデータバスおよびドライブデータバスの構成にしたことにより、内部バスの利用率が低減されて内部バスの転送速度の向上がはかられ、その結果安価で、且つデータ転送が高速化されたディスクアレイ装置を実現することができる。

【0015】また、前記構成により、ディスクキャッシュ上の領域とディスク装置インターフェイスとの間において複数バスを用いた転送によって、データの集合・分散が可能となる。

【0016】また、前記構成により、ディスクキャッシュ上の領域とパリティ生成回路との間において複数バスを用いた転送によって、ディスクキャッシュに格納しているデータを用いて冗長データ生成(パリティデータ生成)が可能となる。また、冗長データ生成回路(パリティデータ生成回路)はデータ転送制御手段(DMAC)

に直結されており、ディスクキャッシュ上の領域とディスク装置インターフェイス間のデータ転送バスに対応してデータ転送バスのデータをパリティ生成回路にも入力することを可能にして、ディスクキャッシュ上のデータをディスク装置インターフェイスとパリティ生成回路との両方に転送して、生成した冗長データ(パリティデータ)を、ディスクキャッシュに転送することなく直接ディスク装置に転送することができ、冗長データ生成時(パリティデータ生成時)におけるディスクキャッシュへのトラフィックが低減され、大容量で連続した領域に対するライト命令に対してディスクアレイ装置における実効転送速度を向上させることができる。

【0017】

【実施例】本発明の実施例を図1～図5を用いて具体的に説明する。図1は、本発明の一実施例であるホストコンピュータ17とディスクアレイ装置18によるシステムの構成図を示す。ホストコンピュータ17とディスクアレイ装置18は、ホストインターフェイス3で接続されている。ホストインターフェイス3は、1チャンネルで、最大転送速度が20MB/sである2バイトバス幅のSCSI-2である。なお、ホストインターフェイス3は、複数チャンネルで構成しても良い。ホストコンピュータ17は、ホストインターフェイス3を介し、SCSI(Small Computer System Interface: 小型コンピュータ用周辺装置インターフェイスの一種。ANSIで規格されている。)プロトコルに従い、ディスクアレイ装置18をアクセスする。

【0018】ディスクアレイ装置18は、アレイコントローラ1と複数のディスク装置2で構成される。ディスクアレイ装置18は、RAID5等のアーキテクチャを採用し、RAID5冗長度はデータディスク4台に対し、パリティディスク1台であり、ディスク装置5台をRAID5の列構成の単位とする。また、ディスクアレイコントローラ1は、5チャネルのドライブインターフェイス16を有し、複数のディスク装置2を接続している。ドライブインターフェイス16は、最大転送速度が10MB/sである1バイトバス幅のSCSI-2である。

【0019】更に、ディスクアレイ装置18で用いられたディスク装置2の性能を図2に示す。即ち、ディスク装置におけるスピンドル回転数は5400rpm、平均シーク時間は9ms、サステインデータ転送速度は7.2MB/s、SCSIデータ転送速度は10MB/sである。

【0020】次にディスクアレイコントローラ1の構成について説明する。アレイコントローラ1は、MPU8、MPUコントロール回路9、ROM10、ワークRAM11からなるディスクアレイ制御部と、ホストインターフェイスコントローラ4、高速にデータを一時格納するDRAM等で構成されたディスクキャッシュ(disk c

11

ache) 5、データ転送制御手段であるD M A C (Direct Memory Access Controller: 小型コンピュータの中で、C P UやI/Oプロセッサの代わりに、メモリー←→メモリー間またはメモリー←→I/O間のデータ転送を制御する部分。データ転送に必要な、ソース・アドレス、ディスティネーション・アドレスを生成し、ソースの読み出しサイクル、ディスティネーションの書き込みサイクルを駆動する。) 6、パリティ生成回路(冗長データ生成回路) 7、5チャネルのドライブインターフェイスに対応して一つずつのドライブインターフェイスコントローラ12a～12eからなるユーザデータ転送制御部、及びホストデータバス13、ドライブデータバス14、M P Uバス15の3つのバスを有する。M P U8は、R O M10に記憶されたプログラムに基づいてワークR A M11に書き込み、読み出ししながらM P Uコントロール回路9により、M P Uバス15を介してユーザデータ制御部を構成するD M A C6、パリティ生成回路7、及びドライブインターフェイスコントローラ12a～12eを制御する。D M A C6は、ホストデータバス13、ドライブデータバス14上でのユーザデータの転送を制御し、R A I D5特有のデータの集合・分散を行う。パリティ生成回路7は、ユーザデータに対して冗長データであるパリティデータの演算を行い、また1台のディスク装置2が障害時にユーザデータの回復処理を行う。パリティ生成回路7は、D M A C6に直結され、ディスクキャッシュ5とドライブデータバス14との間に存在する。

【0021】ホストデータバス13は、ホストインターフェイスコントローラ4とディスクキャッシュ5との間におけるユーザデータ転送用のバスである。ドライブデータバス14は、ディスクキャッシュ5とドライブインターフェイスコントローラ12a～12eとの間におけるユーザデータ転送用のバスである。M P Uバス15は、ユーザデータ制御部を構成するD M A C6、パリティ生成回路7、及びドライブインターフェイスコントローラ12a～12eを制御する制御用のバスである。

【0022】このように3つのバス13、14、15の構成にしたことにより、ホストインターフェイスコントローラ4及びドライブインターフェイスコントローラ12には、安価で、最大転送速度が20MB/sであるS C S IコントロールL S Iを採用することができた。即ち、3つのバス13、14、15の構成にしたことにより、ホストインターフェイス3とホストデータバス13間、ドライブインターフェイス16とドライブデータバス14間にはS C S IコントロールL S I以外にF I F O (First In First Out)などの専用回路は必要とすることなく、ホスト17とディスクキャッシュ5との間におけるデータ転送と、ディスクキャッシュ5とディスク装置2a～2eとの間におけるデータの転送とを同時に実行することができる。ホスト17、即ちユーザとの間のイン

12

タフェイスであるホストインターフェイス3における2バイトバス幅のS C S I-2の最大転送速度は、規格から20MB/sであるため、ホストデータバス13、ドライブデータバス14の転送速度をそれぞれ20MB/sとなり、ディスクキャッシュ5のデータ転送速度は、ホスト17とディスクキャッシュ5との間における転送と、ディスクキャッシュ5とディスク装置2a～2eとの間における転送とを同時に実行するため、40MB/sとなり、データ転送が十分可能となる。

10 【0023】D M A C6は、チャネルa、チャネルb、チャネルcの3チャネルを有し、3チャネルを使った同時転送が可能である。即ち、チャネルaは、ディスクキャッシュ5-ディスク装置2a～2eの間における転送、ディスクキャッシュ5-パリティ生成回路7とディスクキャッシュ5の間における転送、ディスクキャッシュ5-パリティ生成回路7-ディスク装置2a～2eの間における転送に用いる。チャネルbおよびチャネルcは、ディスクキャッシュ5-ディスク装置用2a～2eの間およびホスト17-ディスクキャッシュ5の間におけるデータ転送専用に用いる。

20 【0024】次に、D M A C6におけるチャネルaの機能及びD M A C6におけるチャネルaによるデータ転送について説明する。即ち、D M A C6内には、D M A C6チャネルaについて5つのキャッシュアドレスカウンタ201とキャッシュアドレスカウンタ201に対応する5つのポート指定レジスタ202とキャッシュアドレスカウンタ201に対応する5つの並行転送指定レジスタ205が備えられている。キャッシュアドレスカウンタ201は、ディスクキャッシュ5上の領域203のアドレスを指定するカウンタである。ポート指定レジスタ202は、キャッシュアドレスカウンタ201に対応するドライブインターフェイス16a～16eのチャネルまたはパリティ生成回路7を指定するレジスタである。キャッシュアドレスカウンタ201とこれに対応するポート指定レジスタ202により、ディスクキャッシュ5-ドライブインターフェイス16a～16eの間、またはディスクキャッシュ5-パリティ生成回路7の間における5系統のバス204を指定することができる。D M A C6におけるチャネルaは、時分割で、指定された複数の転送バス204の定まった単位のデータを順番に転送する機能を持つ。また、並行転送指定レジスタ205は、対応するバス204のデータをパリティ生成回路7にも同時に並行転送することを指定するレジスタである。

30 D M A C6におけるチャネルaは、複数バス204による転送機能に加えて、複数バス204のデータをパリティ生成回路7にも同時に並行転送する機能を持つ。

40 【0025】図3は、D M A C6におけるチャネルaを、ディスクキャッシュ5-ディスク装置2a～2eの間におけるデータ転送に使用する場合を示す。図4は、D M A C6におけるチャネルaを、ディスクキャッシュ

13

5-パリティ生成回路7-ディスクキャッシュ5の間ににおけるデータ転送に使用する場合を示す。図5は、DMAC6におけるチャネルaを、ディスクキャッシュ5-ディスク装置2a～2dの間ににおける転送とディスクキャッシュ5-パリティ生成回路7-ディスク装置2eの間ににおける同時並列転送に使用する場合を示す。

【0026】図3に示すように、MPU8は、MPUバス15を介し、DMAC6のチャネルaにおけるキャッシュアドレスレジスタ201a～201eに、ディスクキャッシュ5の領域203a～203eのアドレスを指定し、キャッシュアドレスレジスタ201a～201eに対応するポート指定レジスタ202a～202eに、ドライブインターフェイス16a～16eのチャネルを指定することで、ディスクキャッシュ5上の領域203a～203eとドライブインターフェイス16a～16eとの間ににおけるバスを5つ指定することができる。また、データ転送を行なうディスク装置2は、SCSIコマンド発行によってあらかじめ選択されており、ドライブインターフェイス16a～16eのチャネルごとにドライブインターフェイス16a～16e-ディスク装置2a～2eのバスが存在している。

【0027】従って、DMAC6において、チャネルaの複数バス204a～204eによる転送機能を用いると、ホスト17からホストデータバス13を用いてディスクキャッシュ5上に格納されたユーザデータを、ストライプサイズと呼ばれる一定の大きさに分割して複数のディスク装置2a～2eに分散転送可能である。また、上記ユーザデータが、ディスクキャッシュ5上の複数の不連続領域に格納されている場合でも、ディスクキャッシュ5上の任意の5の領域203のデータをそれぞれ異なるディスク装置2a～2eに分散転送可能である。

【0028】また、ディスク装置2からホスト17へデータの転送方向を変えることで、この複数バスによる転送機能により、複数のディスク装置2a～2eからのデータを、ディスクキャッシュ5上の連続領域に集合転送することが可能である。

【0029】更に、図4に示すように、DMAC6におけるチャネルaにおいて、ポート指定レジスタ202a～202dにパリティ生成回路7を指定して、ディスクキャッシュ5上の複数の領域203a～203dとパリティ生成回路7との間ににおけるバス204a～204dを複数指定することにより、該指定された複数バス204a～204dに基づく転送機能によりディスクキャッシュ5上の複数の領域203a～203dにおけるユーザデータを用いたパリティ生成回路7におけるパリティデータ演算と、該演算されたパリティデータを、ポート指定レジスタ202eおよびキャッシュアドレスレジスタ201eを介してディスクキャッシュ5の領域203eに転送することが可能となる。

【0030】また、図5に示すように、DMAC6にお

50

14

けるチャネルaは、前述の複数バス204a～204dによるデータ転送に加えて、それぞれのバス204a～204dに対応する平行転送指定レジスタ205a～205dにより、それぞれのバス204a～204dの転送データをパリティ生成回路7に平行入力する機能を持つ。また、5系統のうちの残りのバス204eを用いて、パリティ生成回路7の出力を、キャッシュアドレスレジスタ201e、ポート指定レジスタ202eおよび平行転送指定レジスタ205eを介してディスク装置2eに格納することができ、その結果DMAC6におけるチャネルaにおいて、ディスクキャッシュ5-ディスク装置2a～2eの間にデータ転送と、ディスクキャッシュ5-パリティ生成回路7-ディスク装置2eの間に同時並列データ転送とが可能となる。

【0031】ディスクアレイ装置18では、ホスト17からの大容量で、連続領域に対するリード命令については、DMAC6におけるチャネルaおよびチャネルbを用いて、ディスク装置2a～2e-ディスクキャッシュ5の間ににおけるデータ転送と、ホスト17-ディスクキャッシュ5の間ににおけるデータ転送を同時に実行する。ディスク装置2a～2e-ディスクキャッシュ5の間ににおけるデータ転送には、DMAC6におけるチャネルaの複数バス204a～204eによる転送機能を用いて複数のディスク装置2a～2eからデータをディスクキャッシュ5上の連続領域へ集合転送する。

【0032】また、ディスクアレイ装置18では、ホスト17からの大容量で連続領域に対するライト命令については、DMAC6におけるチャネルaおよびチャネルbを用いて、ディスクキャッシュ5-ディスク装置2a～2eの間ににおけるデータ転送と、ホスト17-ディスクキャッシュ5の間ににおけるデータ転送とを同時に実行する。ディスクキャッシュ5-ディスク装置2a～2dの間ににおけるデータ転送には、DMAC6におけるチャネルaの複数バス204a～204dによる転送機能と、それぞれのバス204a～204dにおける転送データを、パリティ生成回路7に平行入力する機能を用いてディスクキャッシュ5のデータを複数のディスク装置2a～2dと、パリティ生成回路7の両方に同時に転送し、パリティ生成回路7で生成されたパリティデータを、ディスクキャッシュ5を介すことなく直接、バス204eを用いてキャッシュアドレスレジスタ201e、ポート指定レジスタ202eおよび平行転送指定レジスタ205eを介してディスク装置2eに転送する。

【0033】本実施例では、ディスクアレイ制御部を、シングルプロセッサ構成としており、ホストインターフェイスコントローラ4及びドライブインターフェイスコントローラ12には、安価で、最大転送速度が20MB/sである市販のSCSIコントロールLSIを採用し、ホストインターフェイス3とホストデータバス13間、ドライブインターフェイス16とドライブデータバス14間に

はSCSIコントロールLSI以外にFIFOなどの専用回路は存在しないため、ディスクアレイコントローラの原価を、ディスクアレイ制御部がマルチプロセッサ構成(MPU8個)で、MPUバス15とユーザデータ転送バスの2本の内部バス13、14との構成により、1/5に低減でき、その結果ディスクアレイ装置18として低価格化を達成することができる。

【0034】また、ディスクアレイ装置18では、パーティ生成回路で生成されたパーティデータを直接ディスク装置に転送するように構成したので、ホスト17からの大容量で連続領域に対するライト命令の実効転送速度を、パーティをディスクキャッシュ5に一旦記憶する方式に比べて約40%向上させることができる。

【0035】更に、ドライブインターフェイス16として、今後も標準インターフェイスとして伸びるSCSI-2を採用しているため、今後の高性能なディスク装置を接続して、高性能なディスクアレイ装置を構成することが可能となる。

【0036】

【発明の効果】本発明によれば、ユーザデータ転送制御部都ディスクアレイ制御部とを備え、制御用バス(MPUバス)、ホストデータバスおよびドライブデータバスの少なくとも3本のバス構成にしたことにより、ホストインターフェイスコントローラおよびドライブインターフェイスコントローラとして、安価なインターフェイスコントローラLSIの使用を可能にし、その結果ディスクアレイコントローラはもとよりディスクアレイ装置としても、原価を大幅に低減することができる効果を奏する。

【0037】また、本発明によれば、ディスクアレイコントローラの内部バスの転送速度が、ディスクアレイ装置の性能ネックとなることがなく、内部バスの転送速度を向上させることができ、またバス幅の低減も図れることができるので効果を奏する。また、本発明によれば、ディスクアレイコントローラ内において、パーティ生成時に、ディスクキャッシュのトラフィックを低減して、ユーザからの大容量で連続した領域に対するライト命令に対してディスクアレイ装置の実効転送速度を向上できる効果を奏する。

【図面の簡単な説明】

【図1】本発明に係るホストコンピュータを有するディスクアレイ装置の一実施例を示した構成図である。

【図2】本発明に係るディスクアレイ装置に用いられるディスク装置の性能を示す図である。

【図3】本発明に係るDMACを、ディスクキャッシュ-ディスク装置の間におけるデータ転送に使用する場合を説明するための図である。

10 【図4】本発明に係るDMACを、ディスクキャッシュ-パーティ生成回路-ディスクキャッシュの間におけるデータ転送に使用する場合を説明するための図である。

【図5】本発明に係るDMACを、ディスクキャッシュ-ディスク装置の間における転送と、ディスクキャッシュ-パーティ生成回路-ディスク装置の間における同時に並列転送とに使用する場合を説明するための図である。

【符号の説明】

1…ディスクアレイコントローラ、2、2a~2e…ディスク装置

20 3…ホストインターフェイス、4…ホストインターフェイスクントローラ

5…ディスクキャッシュ、6…DMAC(データ転送制御手段)

7…パーティ生成回路、8…MPU、9…MPUコントロール回路

10…ROM、11…ワークRAM

12、12a~12e…ドライブインターフェイスコントローラ

13…ホストデータバス、14…ドライブデータバス

15…MPUバス(制御用バス)

30 16、16a~16e…ドライブインターフェイス

17…ホストコンピュータ、18…ディスクアレイ装置

201、201a~201e…キャッシュアドレスカウンタ

202、202a~202e…ポート指定レジスタ

204、204a~204e…バス

205、205a~205e…平行転送指定レジスタ

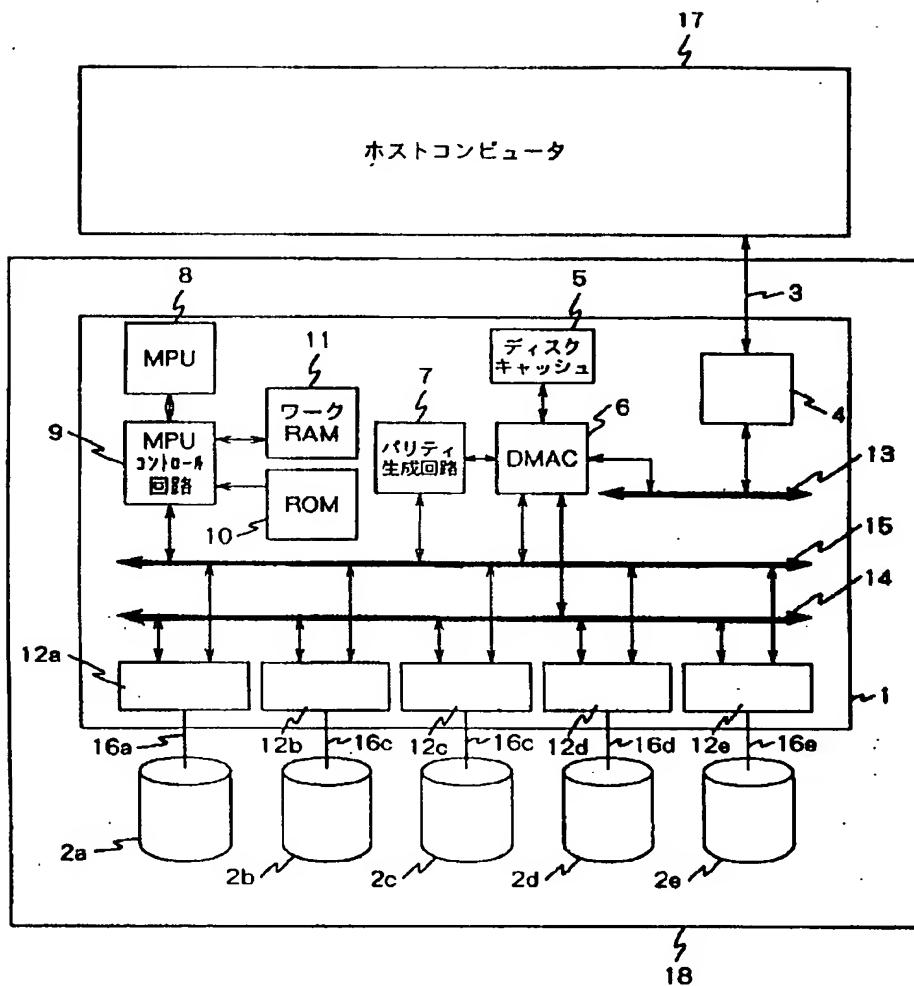
【図2】

図 2

項目	性能
スピンドル回転数 (rpm)	5400
平均シーク時間 (ms)	9
サステインデータ転送速度 (MB/s)	7.2
SCSIデータ転送速度 (MB/s)	10

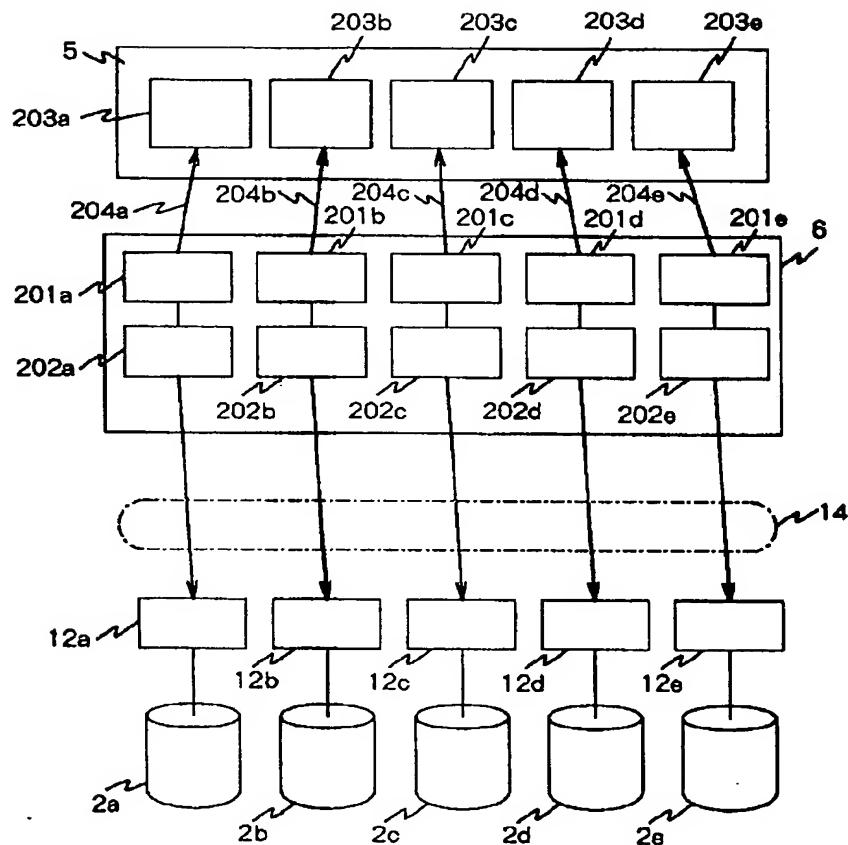
【图1】

1



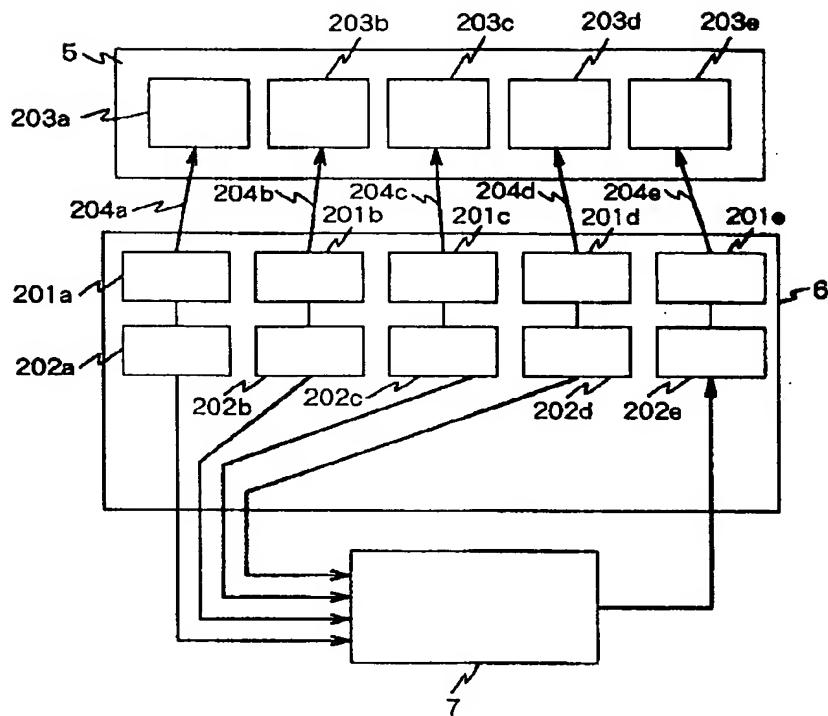
【図3】

図 3



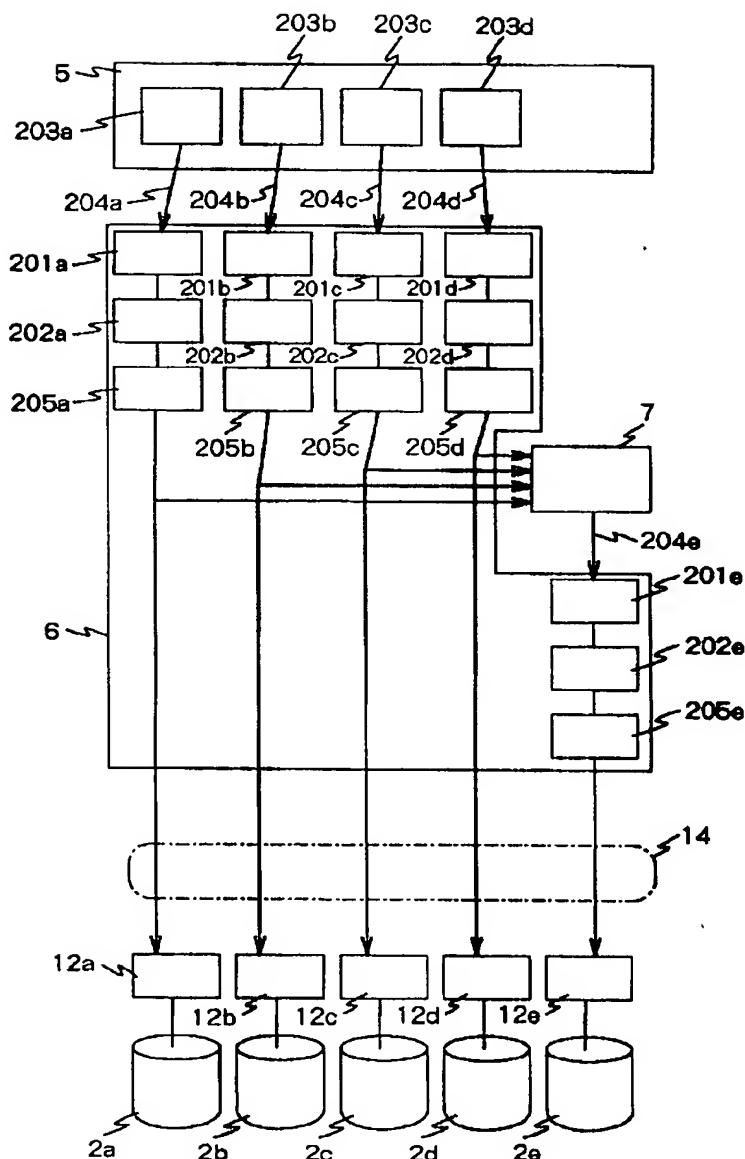
【図4】

図 4



【図5】

図 5



フロントページの続き

(72)発明者 松本 純

神奈川県川崎市麻生区王禅寺1099株式会社
日立製作所システム開発研究所内

(72)発明者 岩崎 秀彦

神奈川県小田原市国府津2880番地株式会社
日立製作所ストレージシステム事業部内